

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 9, 2003, 16:24:23 ; Search time 18.4571 Seconds

(without alignments)  
98.997 Million cell updates/sec

Title: US-09-905-691-2

Perfect score: 19

Sequence: 1 ARAARRAARRAARRAEEA 19

Scoring table: Oligo

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 9616682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

## Post-processing: Listing first 45 summaries

Database : PIR\_76;\*

1: Pir1;\*

2: Pir2;\*

3: Pir3;\*

4: Pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	47.4	105	2 D86976	probable integration host factor [imported] - Mycobacterium leprae
2	9	47.4	190	2 B70899	C; Species: Mycobacterium leprae
3	8	42.1	144	2 H72702	C; Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 @text_change 20-Apr-2001
4	8	42.1	328	2 H83363	R; Cole, S.T.; Eiglmaier, K.; Parkhill, J.; Dutony, S.; Fraser, A.; Hamlin, N.; Holterman, M.A.; Rutherford, K.M.
5	8	42.1	336	2 T50935	Nature 409, 1007-1011, 2001.
6	8	42.1	356	2 A21198	A; Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Title: Massive gene decay in the leprosy bacillus.
7	8	42.1	383	2 F83490	A; Reference number: A86909; PMID:21128732; PMID:11234002
8	8	42.1	710	2 AD3479	A; Accession: D86976
9	8	42.1	897	2 T02808	A; Status: preliminary
10	8	42.1	998	2 T35745	A; Molecule type: DNA
11	7	36.8	69	2 AD2110	A; Residues: 1-105 <STO>
12	7	36.8	71	2 AI3052	A; Cross-references: GB:AU450380; PIDN:CA30048.1; GSDB:GN00147
13	7	36.8	77	2 AH0925	C; Genetics:
14	7	36.8	110	2 T30752	A; Gene: miHF
15	7	36.8	145	2 E75622	RESULT 2
16	7	36.8	157	2 D70772	probable miHF protein - Mycobacterium tuberculosis (strain H37RV)
17	7	36.8	159	2 C83305	C; Species: Mycobacterium tuberculosis
18	7	36.8	174	2 DB7638	C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 @text_change 20-Jun-2000
19	7	36.8	231	2 F75459	R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feilwell, T.; Gentles, S.; Hamlin, N.; Holroyd, R.; Jandaream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
20	7	36.8	252	2 AH3618	Nature 393, 537-544, 1998.
21	7	36.8	263	1 C39741	A; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
22	7	36.8	292	2 JE0233	A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
23	7	36.8	315	2 G87464	A; Reference number: A70500; PMID:9634230
24	7	36.8	326	2 B87557	A; Accession: B70899
25	7	36.8	326	2 S27534	A; Status: preliminary; nucleic acid sequence not shown; translation not shown
26	7	36.8	336	2 E84295	A; Molecule type: DNA
27	7	36.8	338	2 D95851	A; Residues: 1-190 <COL>
28	7	36.8	354	2 B55733	A; Cross-references: GB:Z80108; GB:AL123456; PIDN:93256012; PMID:915428
29	7	36.8	358	2 F87364	A; Experimental source: strain H37RV
					C; Genetics:
					C; Superfamily: Mycobacterium tuberculosis probable miHF protein

Query Match 47.4%; Score 9; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 0.66; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 10 AAARRAAE 18  
Db 104 AAARRAAE 112

RESULT 3  
H72702 hypothetical protein APE1039 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Accession: H72702  
C:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:10382966  
A:Accession: H72702  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-144 <KAN>  
A:Cross references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80024.1; PID:di043810; PID:9510  
A:Experimental source: strain K1  
C:Genetics:

Query Match 42.1%; Score 8; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

Qy 3 AARRAARA 10  
Db 3 AARRAARA 10

RESULT 4  
HB3363 pyoverdine biosynthesis protein PvCA PA2254 [imported] - Pseudomonas aeruginosa (strain  
C:Species: Pseudomonas aeruginosa  
C:Accession: HB3363  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,  
J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A:Reference number: A83950; MUID:10984043  
A:Accession: HB3363  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-328 <STOP>  
A:Cross references: GB:AE004651; GB:AE004091; NID:g9948277; PIDN:AAG05642.1; GSPDB:GN001  
A:Experimental source: strain PAO1  
C:Genetics:

Query Match 42.1%; Score 8; DB 2; Length 328;  
Best Local Similarity 100.0%; Pred. No. 6.6; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

Qy 1 ARARRAA 8  
Db 319 ARARRAA 326

RESULT 5  
T5935 isomerase/decarboxylase homolog Dith [imported] - Pseudomonas abietaniphila  
C:Species: Pseudomonas abietaniphila

C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50935  
R:Martin, V.J.; Mohn, W.W.  
J. Bacteriol. 181, 2675-2682, 1999  
A:Title: A novel aromatic-ring-hydroxylating dioxygenase from the diterpenoid-degradi  
A:Reference number: 222821; MUID:99235742; PMID:10217753  
A:Accession: T50935  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-336 <MAP>  
A:Cross-references: EMBL:AF119621; PIDN:ADD21070.1  
A:Experimental source: strain BRME-9; ATCC700689  
C:Genetics:  
A:Gene: dith

Query Match 42.1%; Score 8; DB 2; Length 326;  
Best Local Similarity 100.0%; Pred. No. 6.8; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

Qy 10 AAARRARA 17  
Db 105 AAARRARA 112

RESULT 6  
A21198 H-2 class I histocompatibility antigen pH-2D-24 - mouse  
C:Species: Mus musculus (house mouse)  
C:Accession: A21198  
R:Alalanne, J.L.; Cochet, M.; Kummer, A.M.; Gachelin, G.; Kourilsky, P.  
Proc. Natl. Acad. Sci. U.S.A. 80, 7561-7565, 1983  
A:Title: Different exon-intron organization at the 5' part of a mouse class I gene is  
A:Reference number: A21198  
A:Accession: A21198  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-356 <LAL>  
A:Cross-references: GB:K01182; NID:g199544; PIDN:AA39653.1; PID:g38468  
A:Note: the authors translated the codon CCC for residue 288 as Ser.  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
<IMMD>  
C:20-27/Domain: immunoglobulin homology

Query Match 42.1%; Score 8; DB 2; Length 356;  
Best Local Similarity 100.0%; Pred. No. 7.1; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0;

Qy 6 RAARRAAR 13  
Db 20 RAARRAAR 27

RESULT 7  
F83490 probable multidrug resistance efflux pump PA1237 [imported] - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C:Accession: F83490  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,  
J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: A83950; MUID:2043737; PMID:10984043  
A:Accession: F83490  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-383 <STOP>  
A:Cross-references: GB:AE004553; GB:AE004091; NID:g9947164; PIDN:AAG04626.1; GSPDB:GN  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA1237  
C:Superfamily: multidrug resistance protein A; lipoyl/biotin-binding homology

Query Match 42.1%; Score 8; DB 2; Length 383;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C;Accession: T35745  
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,  
submitted to the EMBL Data Library, August 1999  
A;Reference number: Z21568  
A;Accession: T35745  
A;Status: preliminary; translated from GB/EMBL/DDJB  
A;Molecule type: DNA  
A;Residues: 1-998 <SAU>  
A;Cross-references: EMBL:ALL109732; PIDN:CAB52056.1; GSPDB:GN00070; SCOEDB:SC7H2.14

Qy 8 ARAARRA 15  
| | | | |  
Db 187 ARAARRA 194

RESULT 8  
AD3479 ATP-dependent helicase hrpB BMEI1818 [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Accession: AD3479  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 14-Apr-2003  
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Mazur, M.; Gotsman, E.; Elzra, P.H.; Hagius, S.; O'callaghan, D.; Letessier, A.; Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A;Reference number: AD3479  
A;Accession: AD3479  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-710 <KUR>  
A;Cross-references: GB:AE008917; PIDN:AAL52999.1; PID:q17983853; GSPDB:GN00190  
C;Genetics:

A;Gene: BMEI1818  
A;Map position: I  
C;Superfamily: HrpB type ATP-dependent RNA helicase

Query Match 42.1%; Score 8; DB 2; Length 998;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
C;Species: Nostoc sp. PCC 7120  
C;Accession: AD2110  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquio, Nakazaki, N.; Shimpoo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, DNA Res. B, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Nostoc sp. strain PCC 7120  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AD2110  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-69 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BAR74134.1; PID:q17131527; GSPDB:GN00179  
C;Genetics:

A;Gene: asl2435

RESULT 9  
T02808 conserved hypothetical protein YPL199c, L2602.6 [imported] - Leishmania major (strain F)  
C;Species: Leishmania major  
C;Accession: D81457; T02808  
R;Myler, P.J.; Audley, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; Proc. Natl. Acad. Sci. U.S.A., 96, 2902-2906, 1999  
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes  
A;Reference number: A81455; MUID:99178897; PMID:10077609  
A;Accession: D81457  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-897 <PYL>  
A;Cross-references: GB:AE001274; NID:93264850; PIDN: AAC24631.1; PID:q92995584; GSPDB:GN00000000  
C;Genetics:

A;Gene: L2602.6  
A;Map position: 1

Query Match 42.1%; Score 8; DB 2; Length 897;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
C;Species: Agrobacterium tumefaciens (strain C58, Dupo  
C;Accession: AI3052  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan, P.; Romero, P.; Zhang, S.; Karp, P.; Gordon-Kamm, E.W.; Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AI3052  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-71 <KUR>  
A;Cross-references: GB:AE008689; PIDN:AAI44839.1; PID:q17742483; GSPDB:GN00187  
RESULT 10  
T35745 probable ATP-binding RNA helicase - Streptomyces coelicolor

A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atu4038  
 A;Map position: linear chromosome  
 Query Match 36.8%; Score 7; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 RAARRAA 8  
 Db 39 RAARRAA 45

RESULT 13  
 AH0255 hypothetical Protein SRY3665 [imported] - *Salmonella enterica* subsp. *enterica* serovar Ty  
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
 A;Note: this species has also been called *Salmonella typhi*  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: AH0255  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Reference number: AB0022; PMID:21534947; PMID:11677608  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-77 <PAR>  
 A;Cross references: GB:AL513382; PIDN:CAD09426.1; PID:gi16504543; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: SRY3665

Query Match 36.8%; Score 7; DB 2; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 13 RRARAA 19  
 Db 30 RRARAA 36

RESULT 14  
 T30752 hypothetical protein 150R - *Molluscum contagiosum* virus 1  
 N;Alternate names: MC150R  
 C;Species: *Molluscum contagiosum* virus 1  
 C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jul-2000  
 C;Accession: T30752  
 R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
 Science 273, 813-816, 1996  
 A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
 A;Reference number: 220876; PMID:9632459; PMID:8670425  
 A;Accession: T30752  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-110 <SEN>  
 A;Cross references: EMBL:U60315; NID:gi1491943; PIDN: AAC55278.1; PID:gi1492093  
 C;Genetics:  
 A;Note: MC150R  
 C;Superfamily: *Molluscum contagiosum* virus 1 hypothetical protein 150R  
 Query Match 36.8%; Score 7; DB 2; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 10 AAARRAR 16  
 Db 91 AAARRAR 97